University of Cape Town

Parameter estimation of a bivariate diffusion process: The Heston Model

Author: Siyabulela Nomoyi
Supervisor: Dr Melvin Varughese

Department of Statistical Science

Masters of Philosophy in Mathematics of Finance (Economics department)

University of Cape Town, Cape Town, South Africa

Email: snomoyi@gmail.com Handed in: 18 November 2011

Abstract

The main objective of the research is to estimate the parameters on the Heston (1993) model, which models the movement of asset prices assuming that the asset price volatility is stochastic. The paper concentrates on estimating these parameters by approximating the transitional probabilities of the diffusion process with a saddlepoint distribution. By solving a system of ordinary differential equations that are in terms of the system's cumulants, and using these solutions to calculate the saddlepoint, the transitional probabilities of the diffusion process can be approximated. This can subsequently explore the likelihood of the diffusion process through a MCMC routine and thereby making it possible to estimate the parameters. The method is applied to freely available South African market data and the estimated parameters are used in pricing of options by the Heston model. The results are compared to some well known option pricing models.

Keywords: Saddlepoint Approximation; Stochastic Volatility; Multivariate-Coupled Diffusion; Heston Model; Ordinary Differential Equations; Black-Scholes; Markov Chain Monte Carlo.

1. Introduction

Introduction of the Black-Scholes (1973) model caused a major breakthrough in pricing of options in the derivative market. The option pricing formula can be applied to vanilla options and many more other types of options. As with any other financial model there are assumptions that need to be considered for the model to be explained and evaluated: the model is based on a traded stock with a log normal distribution, the stock pays no dividends and there are no penalties in short selling of the stock where the short term interest rate is known and is constant. These assumptions cause the model to be in terms of the stock price and known constant variables. A major short-coming of the model is the assumption of constant volatility of the underlying asset in the priced options. Evidence of the failure of this assumption can be seen from Stein (1989), Aít-Sahalia and Lo (1998) and other papers. If we calculate a basket of options (vanilla puts or calls) and plot the implied volatility against different strike prices and different exercise dates of the options, instead of getting a constant horizontal surface (constant volatility), the graph will have what is called a volatility smile, see Backus et al (1997).

Due to the unrealistic assumption of constant volatility in the Black-Scholes (1973) model, different methods of pricing options needed to be introduced, especially ones that would avoid constant volatility and rather have stochastic volatility. These models account for random change in the volatility of the stock or the asset in question, in contrast to the constant volatility assumption made in the Black-Scholes model. Some models used existing interest rate models to model the instantaneous volatility, like the Stein and Stein (1991) model which has the instantaneous volatility modeled as an Ornstein-Uhlenbeck process. Heston (1993) also follows this approach, with the Cox-Ingersoll-Ross (CIR) as the borrowed interest rate model of the volatility. Though Heston states that the Black-Scholes model can be successful in explaining the stock option prices, it performs badly when used to price foreign currency options, see Melino and Turnbull (1990, 1991). Gkamas and Paxson (1999) tested several option pricing models that incorporate stochastic volatility and found that they remove a large proportion of the bias in the option prices that is usually found in the Black-Scholes, but the models do not completely remove the bias.

The rest of the paper is organized as follows. In Section 2, diffusion processes are introduced and in section 3, the application of the saddlepoint approximation method is covered and generalized for both the univariate diffusion process case and the multivariate diffusion process case. In section 4 the Heston model is introduced as it is the model of interest in the research. Section 5 introduces the Ordinary Differential Equations (ODE) for the cumulants of the Heston model and the MCMC algorithm to be used in the approximation of the parameters of the Heston model. Section 6 introduces the market data used and discusses the results of the parameter estimations. Section 7 compares option pricing models in the collected market data using the estimated parameters. Section 8 discusses some conclusions of the study.

2. Diffusion processes

Looking at the one dimensional case, a diffusion process $\{S(t)\}$ is a solution of the form:

$$S(t) = s_0 + \int_0^t \mu\big(k, S(k)\big)dk + \int_0^t \sigma\big(k, S(k)\big)dW(k) \tag{2.1}$$

where $dW(k)$ is a standard Brownian motion. In equation (2.1), $\mu\big(k, S(k)\big)$ is referred to as the drift of the diffusion process and $\sigma\big(k, S(k)\big)$ is the diffusion function. Both functions are considered to be continuous both in time t and in $S(t)$, such that

$$\int_0^t \sigma^2\big(k, S(k)\big)dW(k) < \infty.$$

An n-dimensional diffusion process is defined as

$$d\boldsymbol{\pi}_t^* = \boldsymbol{\mu}(\boldsymbol{\pi}_t^*; \boldsymbol{\phi})dt + \boldsymbol{\sigma}(\boldsymbol{\pi}_t^* t; \boldsymbol{\phi})d\boldsymbol{W}_t, \tag{2.2}$$

where $\boldsymbol{\pi}_t^* = (\pi_i)_{i=1,2,\ldots,n}$ is a vector of variables that describe the behavior of the diffusion process with time, $\{W_t\}_{t\geq 0}$ is an n-dimensional Weiner process, and $\boldsymbol{\phi} = (\phi_i)_{i=1,2,\ldots p}$ is the vector of parameters of the model and is what is estimated in this project. $\boldsymbol{\mu}(\boldsymbol{\pi}_t^* t; \boldsymbol{\phi}) = (\mu_i)_{i=1,2,\ldots,n}$ is the drift vector of the stochastic process and $\boldsymbol{\sigma}^2(\boldsymbol{\pi}_t^*, t; \boldsymbol{\phi}) = (\sigma_{ij})_{i,j=1,2,\ldots n}$ is the covariance matrix of the process.

Varughese (2010) points out a very important property of the diffusion process which will lead to the approximation of the saddlepoint being possible. Diffusion processes possess a multi-normal distribution over infinitesimal time intervals, and since diffusion processes are Markovian, this makes it possible to derive the likelihood of continuously sampled diffusion points. Since the data being used occurs in discrete times an alternative approach is required to derive the likelihood. One way forward is to look at the evolution of the probability density of the process. The probability density function of the system in equation (2.2) is given as:

$$\frac{\partial p(\boldsymbol{\pi}_t^*)}{\partial t} = -\sum_{i=1}^{n} \frac{\partial}{\partial \pi_i}[\mu(\boldsymbol{\pi}_t^*, t; \boldsymbol{\theta}) p(\boldsymbol{\pi}_t^*)] + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial \pi_i \, \partial \pi_j}\left[\sigma_{ij}(\boldsymbol{\pi}_t^*, t; \boldsymbol{\theta}) p(\boldsymbol{\pi}_t^*)\right] \tag{2.3}$$

Definition 2.1

A diffusion process is said to be reducible if there exists a transformation which can transform the original diffusion process into a unit diffusion whose volatility is an identity matrix.

To approximate the evolution of these probabilities, Ait-Sahalia (2002) used the Hermite polynomial expansion which is a closed-form function. However, the Hermite expansion approximation is only applicable for reducible diffusion processes. Hence, even though it is superior to many competing methods, it cannot be used in non-reducible multivariate diffusion processes without a corresponding loss of accuracy. Another way to estimate these transitional probabilities is to use a saddlepoint approximation which is introduced in the following section.

3. Saddlepoint Approximation

Daniels (1954) first introduced the saddlepoint approximations to statistics and the computations (which required super computers at that time) were only possible when cheap computing was available. Huzurbazar (1999) mentioned that saddlepoint approximations are very accurate in the tails of a distribution and also give accurate approximations to higher orders. In terms of the saddlepoint approximation, two forms are considered: the univariate and the multivariate form. Models like the Cox-Ingersoll-Ross can be analysed in the univariate case, and a way of approximating the model using the saddlepoint can be seen in Varughese (2010). The applicable case for this project is the multivariate form as the Heston model is a bivariate diffusion process.

The saddlepoint approximation method has been demonstrated to be highly accurate and fast (Varughese, 2010). In this paper, the saddlepoint approximation will be used to predict the Heston model's transitional distribution through time, as Varughese and Fatti (2008) showed that the method can be applied to multivariate diffusion processes. In order to perform the approximation, the moment generating function (MGF) of the process is required. Let the parameter vector of the MGF be given by $\boldsymbol{\Lambda} = (v_1, v_2, \dots v_n)$, and let the partial differentials of the parameters be given by $\boldsymbol{\Lambda}^* = (\partial/\partial v_1, \partial/\partial v_2, \dots, \partial/\partial v_n)$. The MGF of the diffusion process in equation (2.2) is given as

$$M(\boldsymbol{\Lambda}, t) = E[exp(v_1\pi_1(t) + v_2\pi_2(t) + \dots + v_n\pi_n(t))], \qquad (3.1)$$

The cumulants of a system are defined as follows:

$$\kappa_{r_1,r_2,\dots r_n}(t) = \begin{cases} E_t(\pi_i) & for\ r_i = 1, r_j = 0; j \neq i \\ E_t[\prod_{i=1}^n(\pi_i - E_t(\pi_i))^{r_i}] & for\ 1 < \sum_{j=1}^n r_j < 4 \end{cases},$$

where π_i is the ith variable in the system. The cumulants give the cumulant generating function $K(\boldsymbol{\Lambda}, t)$ (**CGF**) of the system as

$$K(\boldsymbol{\Lambda}, t) = \sum_{r_1 \geq 0}\sum_{r_2 \geq 0} \dots \sum_{r_n \geq 0} \frac{(\prod_{i=1}^n v_i^{r_i})\ \kappa_{r_1,r_2,\dots r_n}(t)}{\prod_{i=1}^n r_i!}, \qquad (3.2)$$

The **CGF** can also be expressed as the logarithm of the **MGF** such that

$$K(\boldsymbol{\Lambda}, t) = ln(M(\boldsymbol{\Lambda}, t)) \qquad (3.3)$$

Differentiating equation (3.3) in terms of time yields the following relationship

$$\frac{1}{M}\frac{\partial M}{\partial t} = \frac{\partial K}{\partial t}, \qquad (3.4)$$

and from Varughese (2010), the corresponding partial differential equation (PDE) for the MGF is given as

$$\frac{\partial M(\Lambda, t)}{\partial t} = \left[\sum_{i=1}^n \nu_i \mu_i(\Lambda^*, t) + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \nu_i \nu_j \sigma_{ij}(\Lambda^*, t)\right] M(\Lambda, t). \qquad (3.5)$$

In this paper, it would be interesting to look at the mean, the variance and the skewness of the system and possibly the kurtosis too. This depends on the truncation of the series expansion of the cumulants in equation (3.2). Using the relationship in equation (3.4) and matching coefficients of the left and the right hand side of the system of equations that will be obtained from equation (3.5), a system of ODEs will be obtained and will be much easier to evaluate computationally. The method can be applied to both univariate and multivariate diffusion processes. The following subsections explain the use of the method both in the univariate and the multivariate case.

3.1 *Univariate case*

From Varughese (2010), the saddlepoint approximation function is given as follows:

$$f_m(x) = \left(2\pi \sum_{i=0}^{m-2} \frac{\kappa_{i+2}\theta_0^i}{i!}\right)^{-\frac{1}{2}} exp\left(\sum_{i=1}^{m} \frac{\kappa_i\theta_0^i}{i!} - \theta_0 x\right), \quad (3.5)$$

where $x = \sum_{i=0}^{m-1} \frac{\kappa_{i+1}\theta_0^i}{i!}$, m is the truncation level and θ_0 is a function of x. Taking x as the data entry at a certain point, and κ_i as the i^{th} cumulant, and solving for θ_0 using its relationship with x and substituting into equation (3.5) then the approximation of the saddlepoint can be obtained for that certain point. Equation (2.3) can be used to approximate the evolution of the system's transitional probabilities $p(\pi_{t_i}^*|\pi_{t_{i-1}}^*)$ once the diffusion process' cumulants are obtained. These transitional probabilities can be used to approximate the maximum likelihood of the diffusion process. Using the Markovian property of diffusion processes, the likelihood function can be approximated as:

$$L(\theta) = p(\pi_{t_1}^*) \prod_{i=2}^{N} p(\pi_{t_i}^*|\pi_{t_{i-1}}^*), \quad (3.6)$$

for discrete time points $t_1, t_2, \ldots, t_N$. N is the number of of data points and for large N, $p(\pi_{t_1}^*)$ which is the stationary distribution is ignored, see Ait-Sahalia (2002) and Varughese (2009). Note that equation (3.5) is applied to the univariate models like the CIR model and the Ornstein–Uhlenbeck model. As this paper is mainly concentrated on the Heston model, there is need to introduce the saddlepoint approximation for multivariate diffusion processes.

3.2 *Multivariate case*

If the model has p variables, then from Renshaw (2000) the saddlepoint approximation becomes:

$$f(x_1, x_2, \ldots, x_p) \cong \frac{exp\{K(\nu_1,\nu_2,\ldots,\nu_p) - \nu_1 x_1 - \nu_2 x_2 - \cdots - \nu_p x_p\}}{(2\pi)^{p/2}\sqrt{|K''(\nu_1,\nu_2,\ldots,\nu_p)|}}, \quad (3.7)$$

where K is the cumulant generating function and the x_i's are the data point vectors. Working with a bivariate diffusion process up to a certain order (3.7), the saddlepoint approximation becomes:

$$f(x_1, x_2) \cong \frac{exp\{K(\nu_1,\nu_2) - \nu_1 x_1 - \nu_2 x_2\}}{2\pi\sqrt{|K''(\nu_1,\nu_2)|}}. \quad (3.8)$$

Equation (3.8) is the saddlepoint approximation function for a bivariate process and $|K''(\nu_1,\nu_2)|$, the determinant of the second derivative of K with respect to (ν_1,ν_2), is replaced by $h_2 g_1 - h_1 g_2$ where:

$$g_1 = \kappa_{20} + \kappa_{30}\nu_1 + \kappa_{21}\nu_2 + \frac{\kappa_{40}\nu_1^2}{2} + \kappa_{31}\nu_1\nu_2 + \frac{\kappa_{22}\nu_2^2}{2},$$

$$g_2 = \kappa_{11} + \kappa_{21}\nu_1 + \kappa_{12}\nu_2^2 + \frac{\kappa_{31}\nu_1^2}{2} + \kappa_{22}\nu_1\nu_2 + \frac{\kappa_{13}\nu_2^2}{2},$$

$$h_1 = \kappa_{11} + \kappa_{21}\nu_1 + \kappa_{12}\nu_2^2 + \frac{\kappa_{31}\nu_1^2}{2} + \kappa_{22}\nu_1\nu_2 + \frac{\kappa_{13}\nu_2^2}{2},$$

$$h_2 = \kappa_{02} + \kappa_{03}\nu_2 + \kappa_{12}\nu_1 + \frac{\kappa_{04}\nu_2^2}{2} + \kappa_{13}\nu_1\nu_2 + \frac{\kappa_{22}\nu_1^2}{2}.$$

From Renshaw (2000) $\nu_1$and ν_2 are obtained by solving simultaneously the following equations:

$$\kappa_{10} + \kappa_{20}\nu_1 + \kappa_{11}\nu_2 + \frac{\kappa_{30}\nu_1^2}{2} + \kappa_{21}\nu_1\nu_2 + \frac{\kappa_{12}\nu_2^2}{2} + \frac{\kappa_{40}\nu_1^3}{6} + \frac{\kappa_{31}\nu_1^2\nu_2}{2} + \frac{\kappa_{22}\nu_2^2\theta_1}{2} + \frac{\kappa_{13}\nu_2^3}{6} - x = 0, \quad (3.9)$$

$$\kappa_{01} + \kappa_{11}\nu_1 + \kappa_{02}\nu_2 + \frac{\kappa_{21}\nu_1^2}{2} + \kappa_{12}\nu_1\nu_2 + \frac{\kappa_{03}\nu_2^2}{2} + \frac{\kappa_{31}\nu_1^3}{6} + \frac{\kappa_{22}\nu_1^2\nu_2}{2} + \frac{\kappa_{13}\nu_2^2\nu_1}{2} + \frac{\kappa_{04}\nu_2^3}{6} - y = 0. \quad (3.10)$$

where for the Heston model, x is the stock price and y is the volatility, off the same dates and which will be introduced in the market data section. Here, to get $\nu_1$and ν_2, obviously the cumulants are needed in order to solve equations (3.9) and (3.10) for the two parameters. A solution to these cumulants is explained in chapters below. When $\nu_1$and ν_2 have been found, the values are substituted in equation (3.8) with the corresponding volatility and stock price values to get the saddlepoint approximation. Section 4 deals with the introduction of the model, and how it is used to approximate a closed form solution for option pricing.

4. The Heston Model

Some of the continuous time stochastic volatility models introduced over the years are the Stein and Stein (1991) model, Hull-White (1987) which assumes independence between the stock price and its volatility. With all these different stochastic volatility models introduced, evaluated and revisited over the years, the Heston (1993) model has gained a lot of interest. A closed form solution for the model has not been found, but approximations have been proposed. Franz (2010) mentions that when choosing a model, you need to look at the intuition of the model, its convenience and its tractability. He goes on and says that the

Heston model is regarded as intuitive and convenient and easy to compute numerically and as the model is *affine*, the model is analytically tractable, see Duffie et al (2000).

Definition 4.1: Affine Diffusion process

X(t) is affine if the characteristic function of X(T) is exponential affine in X(t) for all t ≤T. That is, if there exist functions $\varphi(t,u)$ and $\gamma(t,u)$ such that

$$\mathrm{E}\left[e^{uX(T)}|\mathcal{F}_t\right] = e^{\varphi(T-t,u)+\,\gamma(T-t,u)\,X(t)}$$

where $\mathcal{F}_t$ is the filtration of the diffusion process X(t) and for all $u \in i\mathbb{R}$.

In his 1993 paper, Heston introduced a technique of getting a closed form approximation of a European call option with the underlying asset having stochastic volatility. The model allows for correlation between the stochastic volatility and the stock returns, and uses characteristic functions for the approximation of the closed form solution for the Heston model European call option price. Ben (2007) mentions that the Heston model is very popular and has become the industrial standard in the pricing of exotic options. This popularity has come from the Heston model being able to price the European options efficiently using the Fast Fourier Transform (FFT) algorithm, see Zhylyevskyy (2005). The FFT is one of three commonly used methods that can be used to approximate the pricing of a call option for the Heston model. The other two solutions are to get finite difference solutions of the corresponding Partial Differential Equations (PDE) in the model as seen in Brennan and Schwartz (1977). The third method is Monte Carlo simulation as mentioned by Bondarenko et al (2003) combined with variance reduction.

A number of approaches to estimating the Heston model have emerged. Zirilli et al (2007) uses nonlinear filtering with assets and option prices to evaluate the Heston model using maximum likelihood estimation. The filtering process makes use of the stock log returns and the stock's volatility. This is used to calculate the European call option with a specified strike price, risk free rate and maturity. A new call option price is calculated by adding a Gaussian variable of mean zero and a specified standard deviation which depends on the previous call value and contains the Heston model parameters. The new call option price is used to derive the joint probability function which is the solution to the filtering problem. The probabilities contain the Heston model parameters and solving them using maximum likelihood, the parameters are estimated. Parameter estimation of the Heston model can also be achieved by making use of the saddlepoint approximation and unlike Zirilli's method, it does not need market option prices for the estimation of the parameters. The model uses a system of Ordinary Differential Equations (ODE) that are in terms of cumulants of the estimated model up to a certain order, which when solved can be used to calculate the saddlepoint approximation to determine the transitional probabilities of a diffusion process.

Below is the system of differential equations, which show the dynamics of the stock price and the stochastic volatility. This is the Heston (1993) stochastic volatility model:

$$dS_t = rS_t dt + \sqrt{V_t} S_t dW_t^S, \tag{4.1}$$

$$dV_t = \delta(\theta - V_t)dt + \sigma_v \sqrt{V_t} dW_t^V, \tag{4.2}$$

where

- $\{S_t\}_{t\geq 0}$ is the non-dividend paying stock price process,
- r is the stock price drift,
- $\{V_t\}_{t\geq 0}$ is the stochastic volatility process,
- θ is the long-term mean of the volatility process
- δ is the rate of reversion of the volatility process
- $V_t$is refered to as the instantaneous volatility at time t, which in this case is random and has a volatility σ_v. Thus σ_v is referred to as the volatility of the volatility,
- $\{W_t^S\}_{t\geq 0}$ and $\{W_t^V\}_{t\geq 0}$ are Weiner processes with correlation $\rho \in [-1,1]$.

<u>*Proposition 4.1: The Cox-Ingersoll-Ross (CIR)condition*</u>

Assuming that V(0) > 0. If $2\delta\theta \geq \sigma_v^2$ then the volatility process V(t) can never reach 0 and if $2\delta\theta < \sigma_v^2$ then 0 can be reached.

To keep the volatility process positive at all times, the CIR condition is taken into consideration when estimating the parameters. The data points for the stock price and the volatility process are taken at discrete times. Heston (1993) derived a closed form approximation to the price of the European Call option using a model which can be evaluated analytically but with some parts needing some numerical integration. The derivation of the closed-form approximation can be done using methods from Gatheral

(2006) and Riccati (see Reid, 1972) equations. This method is described clearly in Mikhailov and Nögel (2003).

<u>Lemma 4.1: (Ito's Lemma)</u>

If X(t) is a diffusion process and can be written as follows:

$$dX_t = r(X_t, t)dt + \sigma(X_t, t)dW$$

then if f(X(t),t) is a continuous function of X(t), the derivative of the function is of the form:

$$df(X_t, t) = \left\{ r\frac{\partial f}{\partial x} + \frac{\partial f}{\partial t} + \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x^2} \right\} dt + \sigma \frac{\partial f}{\partial x} dW$$

Gatheral uses lemma 4.1 and shows that the valuation equation for the process is given as:

$$\frac{\partial \Pi}{\partial t} + \frac{1}{2}S^2\frac{\partial^2 \Pi}{\partial S^2} + \rho\sigma_v V_t S_t \frac{\partial^2 \Pi}{\partial V \partial S} + \frac{\sigma_v{}^2}{2}V_t\frac{\partial^2 \Pi}{\partial V^2} + rS_t\frac{\partial \Pi}{\partial S} - rV_t = -\delta(\theta - V_t)\frac{\partial \Pi}{\partial V}, \qquad (4.3)$$

where $\Pi(V_t, S_t, t)$ is the price process of the evaluated portfolio, which in this case is the portfolio that consists of the call option and the stock. Now consider a call option $C(t, S_t, K)$on the above stock, expiring at time T with strike K. Let $X(t) = ln\frac{F(t,T)}{K}$, where $F(t, T)$ is the forward price of the stock S_t. The price of the portfolio can now be set equal to the discounted call option value and be expressed as

$$\Pi(V_t, S_t, t) = e^{-r(T-t)}C(t, S_t, K). \qquad (4.4)$$

From Duffie et al (2000) the approximated value of the option is given as

$$C(t, S_t, K) = K(e^{X(t)}P_2(\tau, X(t), V_t) - P_1(\tau, X(t), V_t)), \qquad (4.5)$$

where $\tau = T - t$ and P_2 and P_1 are probabilities. The probabilities are given in terms of their characteristic functions and can be evaluated as follows according to Heston (1993):

$$P_i(X, V, T; \vartheta) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} Re\left[\frac{e^{i\vartheta}f_i(X,V,T;\vartheta)}{i\vartheta}\right] d\vartheta, \qquad (4.6)$$

for i = 1,2.

Heston (1993) shows how to determine the probabilities which can be expressed in terms of their characteristic functions as in equation (4.6), and once the characteristic functions are inverted using a Fourier transform, then the probabilities can be evaluated.

In evaluating the probabilities in equation (4.5), one encounters an integral that is of complex value as in equation (4.6), which can be evaluated using Simpson's rule (Moodley, 2005) and another method that can be used is the Gaussian Quadrature Rule (Moodley, 2005) which computes the integral numerically.

Amongst other researchers; Glass (2007) and Karlsson (2009) mention that the main purpose why the Heston model is a good choice for pricing options in the stochastic volatility space is the fact that the model allows for approximately fitting implied volatility smiles and skews as well.

5. Parameter estimation for the Heston model

5.1 The Ordinary Differential Equations (ODE)

Using equations (4.1) and (4.2), the variance-covariance and the drifts for the system are listed as follows:

$$\mu_1(S_t) = rS_t \qquad \sigma_{11} = (\sqrt{V_t}S_t)^2 = {S_t}^2V_t$$
$$\mu_2(V_t) = \delta(\theta - V_t) \qquad \sigma_{22} = (\sigma_v\sqrt{V_t})^2 = {\sigma_v}^2V_t$$
$$\sigma_{12} = \sigma_{21} = \rho\sqrt{V_t}S_t\sigma_v\sqrt{V_t} = \rho\sigma_vS_tV_t$$

For the Heston model n = 2 in equation (2.3) and the first derivative of the moment generating function in equation (3.5) is as follows:

$$\frac{\partial M(\Lambda,t)}{\partial t} = \left[\nu_1\mu_1(\Lambda^*,t) + \nu_2\mu_2(\Lambda^*,t) + \frac{1}{2}\nu_1^2\sigma_{11}(\Lambda^*,t) + \frac{1}{2}\nu_2^2\sigma_{22}(\Lambda^*,t) + \nu_1\nu_2\sigma_{12}(\Lambda^*,t)\right]M.$$

Using equation (3.5) the partial differential for the MGF is:

$$\frac{\partial M(\Lambda,t)}{\partial t} = \nu_1 r\frac{\partial M}{\partial \nu_1} + \nu_2 M\delta(\theta) - \delta\nu_2\frac{\partial M}{\partial \nu_2} + \frac{1}{2}\nu_1^2\frac{\partial^3 M}{\partial \nu_1^2\partial \nu_2}$$
$$+\frac{1}{2}\nu_2^2{\sigma_v}^2\frac{\partial M}{\partial \nu_2} + \nu_1\nu_2\rho\sigma_v\frac{\partial^2 M}{\partial \nu_1\partial \nu_2}. \qquad (5.1)$$

Using equation (3.4) the following relationships between the partial differential equations of the MGF and the CGF can be derived:

$$\frac{1}{M}\frac{\partial M}{\partial \nu_i} = \frac{\partial K}{\partial \nu_i}; \qquad \frac{1}{M}\frac{\partial^2 M}{\partial \nu_1\partial \nu_2} = \frac{\partial K}{\partial \nu_1}\frac{\partial K}{\partial \nu_2} + \frac{\partial^2 K}{\partial \nu_1\partial \nu_2};$$

$$\frac{1}{M}\frac{\partial^3 M}{\partial \nu_1^2\partial \nu_2} = \left(\frac{\partial K}{\partial \nu_1}\right)^2\frac{\partial K}{\partial \nu_2} + \frac{\partial^2 K}{\partial \nu_1^2}\frac{\partial K}{\partial \nu_2} + 2\frac{\partial^2 K}{\partial \nu_1\partial \nu_2}\frac{\partial K}{\partial \nu_1} + \frac{\partial^3 K}{\partial \nu_1^2\partial \nu_2};$$

Substituting these into equation (5.1), the following expression results:

$$\frac{\partial K}{\partial t}=(v_1 r)\frac{\partial K}{\partial v_1}+v_2\delta(\theta)-\delta v_2\frac{\partial K}{\partial v_2}+\frac{1}{2}v_1^2\left[\left(\frac{\partial K}{\partial v_1}\right)^2\frac{\partial K}{\partial v_2}+\frac{\partial^2 K}{\partial v_1^2}\frac{\partial K}{\partial v_2}+2\frac{\partial^2 K}{\partial v_1\partial v_2}\frac{\partial K}{\partial v_1}+\frac{\partial^3 K}{\partial v_1^2\partial v_2}\right]+$$
$$\frac{1}{2}v_2^2\sigma_v^2\frac{\partial K}{\partial v_1}+v_1v_2\sigma_v\left[\frac{\partial K}{\partial v_1}\frac{\partial K}{\partial v_2}+\frac{\partial^2 K}{\partial v_1\partial v_2}\right] \tag{5.2}$$

A 4th-order expansion of $K(v_1, v_2; t)$ is given by:

$$K(v_1,v_2;t)=\ 1+\ v_1\kappa_{10}(t)+v_2\kappa_{01}(t)+v_1v_2\kappa_{11}(t)+\frac{v_1^2}{2}\kappa_{20}(t)+\frac{v_2^2}{2}\kappa_{02}(t)+$$
$$\frac{v_1v_2^2}{2}\kappa_{12}(t)+\frac{v_1^2v_2}{2}\kappa_{21}(t)+\frac{v_1^3}{3!}\kappa_{30}(t)+\frac{v_2^3}{3!}\kappa_{03}(t)+\frac{v_1v_2^3}{1!3!}\kappa_{13}(t)+\frac{v_1^2v_2^2}{2!2!}\kappa_{22}(t)+\frac{v_2v_1^3}{1!3!}\kappa_{31}(t)+$$
$$\frac{v_1^4}{4!}\kappa_{40}(t)+\frac{v_2^4}{4!}\kappa_{04}(t).$$

Differentiating the above expansion appropriately in terms of $\frac{\partial K}{\partial v_1}$, $\frac{\partial K}{\partial v_2}$, $\frac{\partial^2 K}{\partial v_1^2}$, $\frac{\partial^2 K}{\partial v_1\partial v_2}$, and $\frac{\partial^3 K}{\partial v_1^2\partial v_2}$ leads to the following system of equations:

$$\frac{\partial K}{\partial v_1}=\kappa_{10}(t)+v_2\kappa_{11}(t)+v_1\kappa_{20}(t)+\frac{v_2^2}{2}\kappa_{12}(t)+v_1v_2\kappa_{21}(t)+\frac{v_1^2}{2}\kappa_{30}(t)+\ \frac{v_2^3}{3!}\kappa_{13}(t)+$$
$$\frac{v_1v_2^2}{2!}\kappa_{22}(t)+\frac{v_2v_1^2}{2!}\kappa_{31}(t)+\frac{v_1^3}{3!}\kappa_{40}(t),$$

$$\frac{\partial K}{\partial v_2}=\kappa_{01}(t)+v_1\kappa_{11}(t)+v_2\kappa_{02}(t)+\frac{v_1^2}{2}\kappa_{21}(t)+v_1v_2\kappa_{12}(t)+\frac{v_2^2}{2}\kappa_{03}(t)+\ \frac{v_1v_2^2}{2!}\kappa_{13}(t)+$$
$$\frac{v_2v_1^2}{2!}\kappa_{22}(t)+\frac{v_1^3}{3!}\kappa_{31}(t)+\frac{v_2^3}{3!}\kappa_{04}(t),$$

$$\frac{\partial^2 K}{\partial v_1^2}=\kappa_{02}(t)+v_2\kappa_{21}(t)+v_1\kappa_{30}(t)+\frac{v_2^2}{2!}\kappa_{22}(t)+v_1v_2\kappa_{31}(t)+\frac{v_1^2}{2!}\kappa_{40}(t),$$

$$\frac{\partial^2 K}{\partial v_1\partial v_2}=\kappa_{11}(t)+v_2\kappa_{12}(t)+v_1\kappa_{21}(t)+\frac{v_2^2}{2!}\kappa_{13}(t)+v_1v_2\kappa_{22}(t)+\frac{v_1^2}{2!}\kappa_{31}(t),$$

$$\frac{\partial^3 K}{\partial v_1^2\partial v_2}=\kappa_{21}(t)+v_2\kappa_{22}(t)+v_1\kappa_{31}(t).$$

By substituting these differential equations into equation (5.2) and matching coefficients with the differentials of the expanded $K(v_1, v_2; t)$, the following system of ordinary differential equations is obtained:

$$\kappa_{10}'(t)=r\kappa_{10}(t),$$

$$\kappa_{01}'(t)=\delta\big(\theta-\kappa_{01}(t)\big),$$

$$\kappa_{11}'(t)=-\delta\kappa_{11}(t)+r\kappa_{11}(t)+\rho\sigma_v\big(\kappa_{01}(t)\kappa_{10}(t)+\kappa_{11}(t)\big),$$

$$\kappa_{20}'(t)=\kappa_{01}(t)\kappa_{02}(t)+\kappa_{01}(t)\kappa_{10}^2(t)+2\kappa_{10}(t)\kappa_{11}(t)+\kappa_{21}(t)+2r\kappa_{20}(t),$$

$$\kappa_{02}'(t)=-2\delta\kappa_{02}(t)+\sigma_v^2\kappa_{01}(t),$$

$$\begin{aligned}\kappa_{21}'(t) = {} & \kappa_{02}(t)\kappa_{03}(t) + \kappa_{02}(t)\kappa_{10}^2(t) + 2\kappa_{01}(t)\kappa_{10}(t)\kappa_{11}(t) + 2\kappa_{11}^2(t) + 2\kappa_{10}(t)\kappa_{12}(t) - \delta\kappa_{21}(t) \\ & + \kappa_{10}(t)\kappa_{21}(t) + 2r\kappa_{21}(t) + \rho\sigma_v\big(2\kappa_{10}(t)\kappa_{11}(t) + 2\kappa_{01}(t)\kappa_{20}(t) + 2\kappa_{21}(t)\big),\end{aligned}$$

$$\kappa_{12}'(t) = -2\delta\kappa_{12}(t) + r\kappa_{12}(t) + \rho\sigma_v\big(2\kappa_{02}(t)\kappa_{10}(t) + 2\kappa_{01}(t)\kappa_{11}(t) + 2\kappa_{12}(t)\big) + \sigma_v^2\kappa_{11}(t),$$

$$\begin{aligned}\kappa_{22}'(t) = {} & \kappa_{02}(t)\kappa_{03}(t) + \kappa_{03}(t)\kappa_{10}^2(t) + 4\kappa_{02}(t)\kappa_{10}(t)\kappa_{11}(t) + 2\kappa_{01}(t)\kappa_{11}^2(t) \\ & + 2\kappa_{01}(t)\kappa_{10}(t)\kappa_{12}(t) + 6\kappa_{11}(t)\kappa_{12}(t) + 2\kappa_{10}(t)\kappa_{13}(t) + 2\kappa_{02}(t)\kappa_{21}(t) \\ & - 2\delta\kappa_{22}(t) + \kappa_{01}(t)\kappa_{22}(t) + 2r\kappa_{22}(t) \\ & + 4\rho\sigma_v\left(\kappa_{11}^2(t) + \kappa_{10}(t)\kappa_{12}(t) + \kappa_{02}(t)\kappa_{20}(t) + \kappa_{01}(t)\kappa_{21}(t) + \kappa_{22}(t)\right) \\ & + \sigma_v^2\kappa_{21}(t),\end{aligned}$$

$$\begin{aligned}\kappa_{30}'(t) = {} & 3\kappa_{11}(t)\kappa_{10}^2(t) + 6\kappa_{01}(t)\kappa_{10}(t)\kappa_{20}(t) + 9\kappa_{11}(t)\kappa_{20}(t) + 6\kappa_{10}(t)\kappa_{21}(t) \\ & + 3\kappa_{01}(t)\kappa_{30}(t) \qquad + 3\kappa_{31}(t) + 3r\kappa_{30}(t),\end{aligned}$$

$$\kappa_{30}'(t) = -3\delta\kappa_{03}(t) + 3\sigma_v^2\kappa_{02}(t),$$

$$\begin{aligned}\kappa_{31}'(t) = {} & 6\kappa_{10}(t)\kappa_{11}^2(t) + 3\kappa_{12}(t)\kappa_{10}^2(t) + 6\kappa_{02}(t)\kappa_{10}(t)\kappa_{20}(t) + 6\kappa_{01}(t)\kappa_{11}(t)\kappa_{20}(t) \\ & + 9\kappa_{12}(t)\kappa_{20}(t) + 6\kappa_{01}(t)\kappa_{10}(t)\kappa_{21}(t) + 15\kappa_{11}(t)\kappa_{21}(t) + 6\kappa_{01}(t)\kappa_{22}(t) \\ & + 3\kappa_{02}(t)\kappa_{30}(t) - \delta\kappa_{31}(t) + 3\kappa_{01}(t)\kappa_{31}(t) + 3r\kappa_{31}(t) \\ & + 3\rho\sigma_v\big(2\kappa_{11}(t)\kappa_{20}(t) + \kappa_{10}(t)\kappa_{21}(t) + \kappa_{01}(t)\kappa_{30}(t) + 3\kappa_{31}(t)\big),\end{aligned}$$

$$\begin{aligned}\kappa_{31}'(t) = {} & -3\delta\kappa_{13}(t) + r\kappa_{13}(t) + 3\rho\sigma_v\big(2\kappa_{11}(t)\kappa_{02}(t) + \kappa_{10}(t)\kappa_{03}(t) + \kappa_{01}(t)\kappa_{12}(t) + \kappa_{13}(t)\big) \\ & + 3\sigma_v^2\kappa_{12}(t),\end{aligned}$$

$$\begin{aligned}\kappa_{40}'(t) = {} & 24\kappa_{10}(t)\kappa_{11}(t)\kappa_{20}(t) + 12\kappa_{01}(t)\kappa_{20}^2(t) + 6\kappa_{21}(t)\kappa_{10}^2(t) + 30\kappa_{20}(t)\kappa_{21}(t) \\ & + 12\kappa_{01}(t)\kappa_{10}(t)\kappa_{30}(t) + 24\kappa_{11}(t)\kappa_{30}(t) + 12\kappa_{10}(t)\kappa_{31}(t) + 6\kappa_{01}(t)\kappa_{40}(t) \\ & + 4r\kappa_{40}(t),\end{aligned}$$

$$\kappa_{04}'(t) = -4\delta\kappa_{04}(t) + 6\sigma_v^2\kappa_{03}(t).$$

From the above set of ODEs, it can be seen that the higher the order, the more computation needed to be done, and the more complicated it is to solve the ODEs. Solutions of the ODEs are used in the evaluation of the saddlepoint by substituting them in equations (3.9) and (3.10) in the multivariate case, and in equation (3.5) for the univariate case. By exploring the likelihood of the diffusion process through a Markov Chain Monte Carlo (MCMC) process, this makes it possible to estimate the model parameters. The following subsection introduces the MCMC algorithm that uses the approximation to the transitional probabilities to estimate the parameters.

5.2 The MCMC algorithm

The above ODEs together with the collected data, makes it possible to estimate the parameters of the Heston model; $\delta, \theta, r, \rho, \sigma_v$ in equations (4.1) and (4.2). The system is implemented in Mathematica® v7.0. To prevent the MCMCs from diverging, the conditions of the CIR must be satisfied for this particular model, before accepting a proposed set of parameters for the chains. Once the parameters have been estimated, then appropriate forecasts may be done and option pricing using the estimated parameters can be done.

The estimation procedure follows the MCMC algorithm stated by Varughese (2010). The algorithm is as follows:

1) Approximate the cumulants evolution using the derived ordinary differential equations from section 5.
2) Choose a starting set β_0 of the parameters (appropriate starting values are derived in this case). Set $\beta_{old} = \beta_0$. A method for choosing the starting set $\beta_0$is described in section 6.
3) Propose a jump from the old set of parameter values β_{old} to a new set of parameter values, $\beta_{new} = \beta_{old} + \Delta\beta$ where $\Delta\beta$ is drawn from a suitably chosen proposal distribution.
4) Calculate the likelihoods $L(\beta_{old})$ and $L(\beta_{new})$. The likelihood $L(\beta)$ is calculated as follows:
    i. Set the likelihood, $L(\beta) = 1$
    ii. Set i to 1.
    iii. Use the system of ordinary differential equations from step 1 together with the data values observed at time $t_i = (t_{i_1}, t_{i_2}, \ldots t_{i_N})$ to predict the values of the cumulants at time t_{i+1}.
    iv. Given the data $\pi^*_{t_i}$ at timet_i, the transitional probability distribution at time t_{i+1} can be approximated by a saddlepoint approximation $\vartheta(\pi^*_{t_i}|\pi^*_{t_{i-1}})$ after substituting for the cumulants derived from step iii.
    v. Set the likelihood $L(\beta) = L(\beta) \times p(\pi^*_{t_i}|\pi^*_{t_{i-1}})$.
    vi. Set $i = i + 1$ and if $i < N$ (where N is the number of data points) and then go to step iii. This loops until $i = N$.
5) Accept the proposed parameter values, β_{new} with probability p where:

$$p = \begin{cases} 1 & L(\beta_{new}) \geq L(\beta_{old}) \\ \dfrac{L(\beta_{new})}{L(\beta_{old})} & L(\beta_{new}) < L(\beta_{old}) \end{cases}$$

6) Go back to step 3.

Using this algorithm, MCMC iterations can be ran for both the univariate case and the multivariate case. In step 3, the suitable distribution considered is the normal distribution and the proposed jump will be shown in the analysis of the results. For this jump to be accepted, the new proposed parameters must satisfy the conditions that the correlation parameter is between -1 and 1, all the other four parameters must be strictly greater than zero and the CIR condition must also be satisfied.

In this project, the highest order considered is the 4th. As it will be seen on section 6, the 2nd order case simplifies to the bivariate normal density, which is the analysis that follows right after this section. After using the above algorithm over 700 000 MCMC runs, the first 600 000 runs where trimmed as part of the burn-in period, and is used to plot the histogram of the approximated values. The median is used to approximate the parameters.

6. The Analysis

6.1 The Market Data

Data for the variable $\{S_t\}_{t\geq 0}$ will be the *All Share Index (ALSI)* levels, from the 1st of February 2007 to the 27th of June 2007. For the $\{V_t\}_{t\geq 0}$variable instead of determining the variance from the index levels, the *South African Volatility Index (SAVI)* will be used, running through the same period. The data is extracted from freely available data from DataStream®, and is taken as daily entries.

The SAVI is based on the FTSE/JSE Topi40 index level and it is determined using the at-the-money volatilities, and from the figure below, the correlation between the two variables is easily evident. The observations (S_i, V_i) will be used as the boundary values for running the ordinary differential equations for the Heston model mentioned above, where i runs from 0 to the penultimate pair of data points.

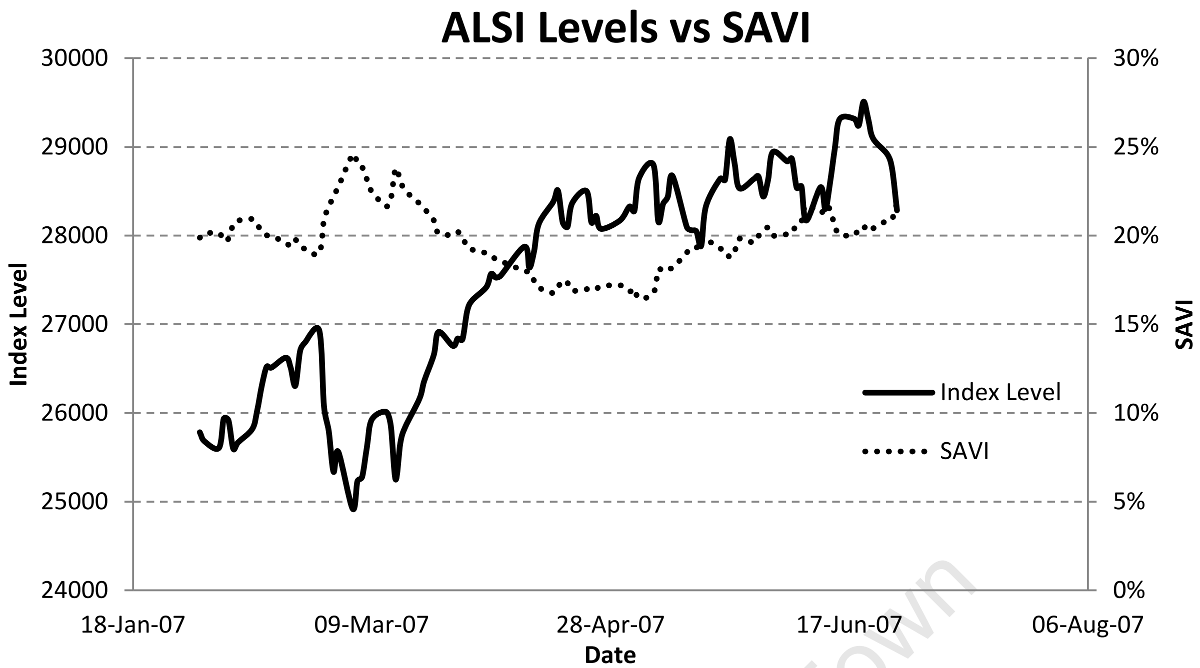

Figure 6.1: The ALSI levels and the movement of the SAVI

In the following subsection, analysis is shown for the 2nd order case by using the bivariate normal density.

6.2 Analysis using a Bivariate Normal

6.2.1 Theory

From equation (3.9), it follows that for *a second order cumulant trancation* the saddlepoint simplifies to a normal distribution. For the multivariate case, as shown by Renshaw (2000) for $n = 2$ the system of equations represented by h_1, g_1, h_2 and g_2 yield the PDF of the bivariate normal (see Appendix A). Hence it is possible to approximate the saddlepoint using the bivariate normal PDF:

$$f(x,y) = \left[2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}\right]^{-1} exp\left\{\frac{-1}{2(1-\rho^2)}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right)\right\} \quad (6.1)$$

Using the above PDF and the cumulants solved by the above ODEs, the following equalities are used:

$x(t) = IndexLevel(t)$; $y(t) = SAVI(t)$; $\mu_1(t) = \kappa_{10}(t)$;$\mu_2(t) = \kappa_{01}(t)$. $\sigma_1^2(t) = \kappa_{20}(t)$; $\sigma_2^2\ (t) = \kappa_{02}(t)$; ρ = *correlation [SAVI, Index level].*

6.2.2 System and Results

Before running the iterations, convergence of the MCMC chains needs to be guaranteed. This is done by carefully choosing the starting values of the parameters so that they are close to the final estimates. The idea is to take into account that for the system to work on the given sets of data, the variance-covariance of the *differenced* data points must be approximately equal to the diffusion matrix of the Heston model. Using the two equations modeling the volatility and the stock, the diffusion matrix of the Heston Model is as follows:

$$\begin{bmatrix} {S_t}^2 V_t & \rho\sigma_v S_t V_t \\ \rho\sigma_v S_t V_t & {\sigma_v}^2 V_t \end{bmatrix}.$$

In order for the system to work properly, a scaling factor for the SAVI dataset is needed to ensure the variance of the differenced indices is the same order of magnitude as predicted by the diffusion matrix. Using the variance of the above system, the idea is to get the following condition holding:

$$Var[dS_t] \cong Mean[{S_t}^2] \times Mean[V_t].$$

Hence the scaling factor k is determined as:

$$k \cong Var[dS_t]/Mean[{S_t}^2].$$

Denoting the scaled SAVI data by V_t^*, the approximations to the starting values for the parameters are shown in table (6.1), taking the time interval to be $t = \frac{1}{252}$.

Table 6.1: Parameter approximated starting values.

Stock price drift	(r)	$Mean[dS_t]$
Long-term mean of the volatility	(θ)	$Mean[d{V_t}^*]$
Rate of reversion	(δ)	$Mean[d{V_t}^*]/(\theta - Mean[{V_t}^*])$
Volatility of the volatility	(σ_v)	$Sqrt(Var[d{V_t}^*]/Mean[{V_t}^*])$
Correlation of Weiner process	(ρ)	$Corr[S_t, {V_t}^*]$

Unlike in Varughese (2010), the true parameters are unknown, so no comparison with true parameter values can be done here. Hence good approximations on this paper rely on how accurate the convergence will be. This is checked by plotting histograms after a certain burn in interval for each parameter, depending on how fast the MCMC chains converge. The proposed parameter *jumps* mentioned in step 3 of section 5 from the old parameter vector to the new is $\{0.02, 0.001, 0.01, 0.003, 0.02\}$ for $\{\rho, \sigma_v, r, \delta, \theta\}$ respectively. For the boundary conditions, starting values for the cumulants $\kappa_{01}(t)$ and $\kappa_{10}(t)$ are set to V_{ti} and S_{ti} respectively, and all of the other cumulants set to zero. As the iterations run through the data, if the current time is t, then $\kappa_{01}(t) = V(t)$, $\kappa_{10}(t) = S(t)$ and all the others are set to 0, and then solving the ODEs using these initial values for the cumulants, the saddlepoint approximation can be obtained. Using the MCMC algorithm for 700 000 iterations with a burn-in of 600 0000, the results for the parameter estimates are shown below.

The algorithm took about 62 hrs to run for the 700 000 iterations, which is approximately 2½ days. During the experiments, it became evident that the time it takes to finish is directly proportional to the number of data values used. Figures 6.2 – 6.4 show the estimates of the parameters' posterior distributions. These are histograms of the sampled values of the MCMC chains after convergence had been achieved. Figure 6.2 for ρ, σ_v and figure 6.3 for r show good signs of convergence, whereas δ and θ show slow convergence and some instability but their values still show convergence.

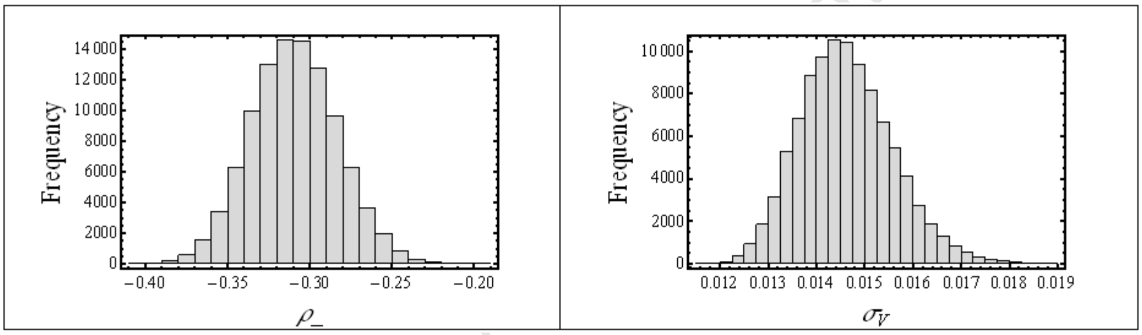

Figure 6.2: The correlation and the vol-of-vol

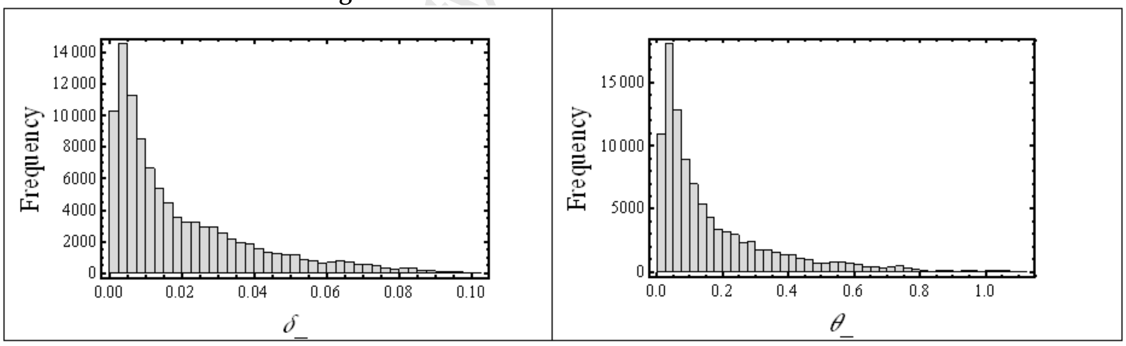

Figure 6.3: The rate of reversion and the mean-reversion

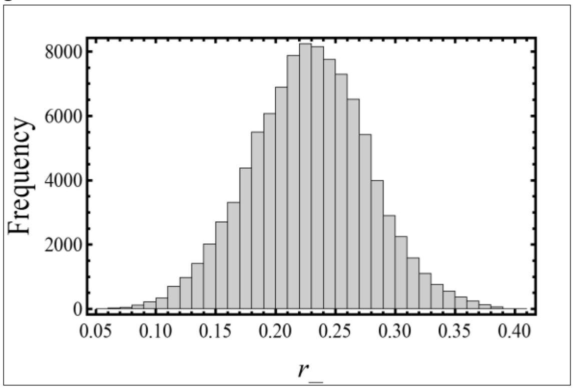

Figure 6.4: The drift

The histograms are drawn using the same burn in periods of 100 000. To estimate the values of the parameter, the median of the parameter values after the burn-in period is used and approximations are given in table (6.2).

Table 6.2: Approximation of the parameters

Parameter		Estimated value
Stock price drift	(r)	0.22892817
Long-term mean of the volatility	(θ)	0.09779974
Rate of reversion	(δ)	0.01197934
Volatility of the volatility	(σ_v)	0.01455406
Correlation of Weiner process	(ρ)	-0.30982197

The graphs below show how fast the convergence was happening. Figure 6.5 below shows the convergence of the mean reversion parameter. The graph shows some instability in the parameter.

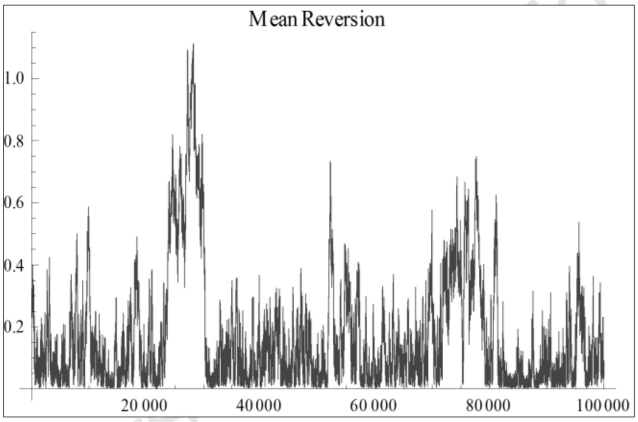

Figure 6.5: Convergence graphs of the mean-reverting parameter

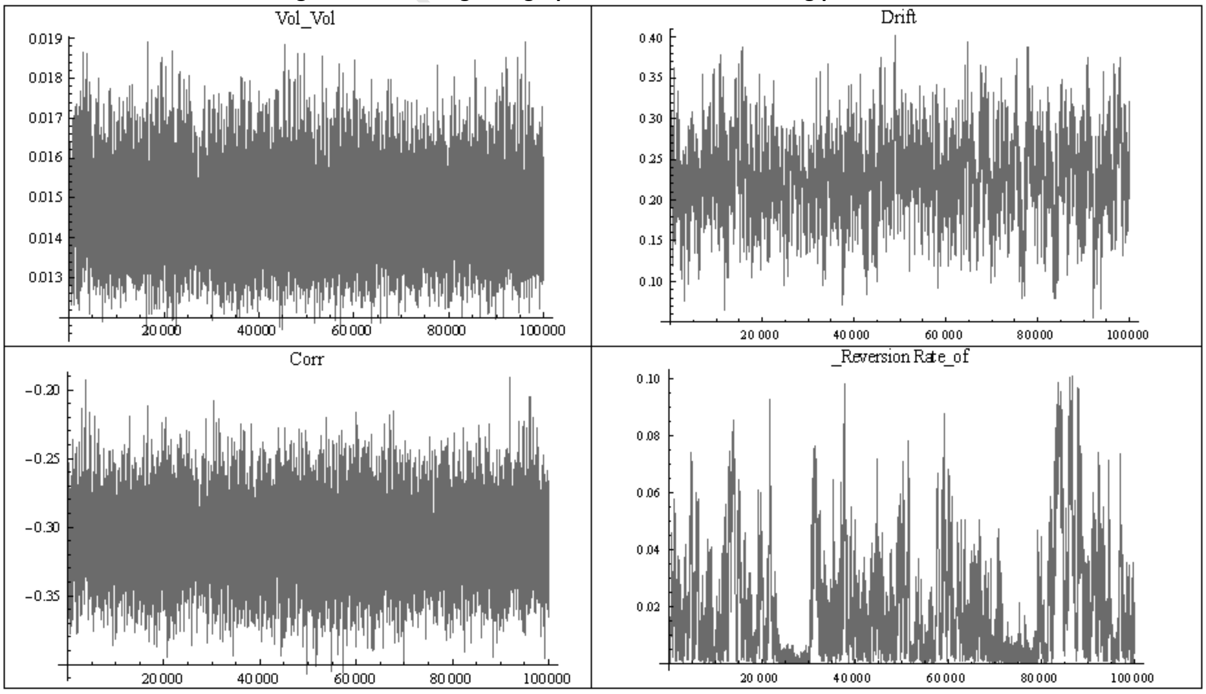

Figure 6.6: Convergence graphs of the parameters.

The Vol-of-Vol parameter, the correlation and the drift parameters show very fast convergence. The rate of reversion parameter also shows some instability as also seen in the mean reversion parameter. The misbehaviour in the reversion parameters can be explored in trying to see if there is some sort of correlation between the two parameters. The correlation between the two parameters is shown in figure 6.7 below.

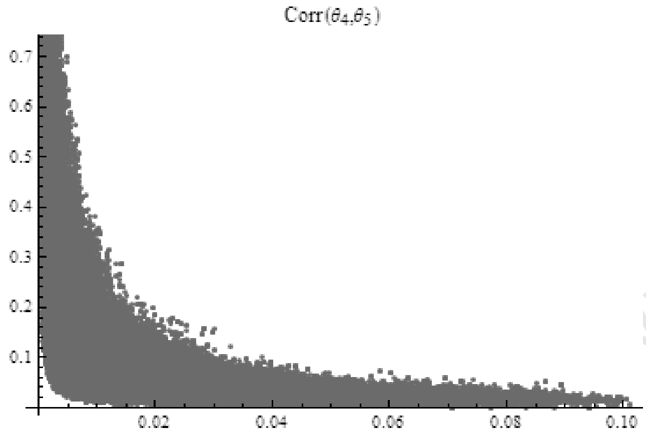

Figure 6.7: Correlation between the mean reversion and the rate of reversion.

From figure 6.7, it can be seen that there is evidence of correlation between the two parameters, which might explain why the parameters behave in the same manner in terms of their convergence.

The saddlepoint approximation seems to be behaving for second order. In the next section the method is checked if it holds for higher orders and discusses some of the problems that came up. The method is also checked at how it behaves using some of the suggestions in Renshaw (2000).

6.3 Analysis using a 4th order cumulant truncation.

6.3.1 Theory

For this part, all the ODEs from section 5.1 are used and the resulting cumalants substituted in equations (3.9), (3.10) and then solving for $\nu_1$and ν_2 to get the saddlepoint approximation at the data points. Mathematica produces four different solutions for the νs, most of which are complex numbers. Choosing the appropriate solution pair (real valued) for the νs yields the saddlepoint approximations. Some of the Real ν values yield a saddlepoint of complex value. Renshaw (2000) shows how to solve for the cumulants for a 4^{th} order CGF and solves the simultaneous equations (3.9) and (3.10) using Newton-Raphson's method. It is difficult to solve these equations, so Renshaw (2000) sets $\kappa_{12}(t)$, $\kappa_{12}(t)$ equal to zero for the 3rd order as well as $\kappa_{31}(t)$, $\kappa_{13}(t)$and $\kappa_{22}(t)$ to zero for the 4th order so as to ensure full support of the saddlepoint. The system is also run without equating the cumulants to zero.

6.3.2 Problems encountered

This method seems to give problems for orders ≥ 3. No problems were encountered when the system was running using the bivariate normal density which represents the 2nd order case. Running the system with cumulants up to the 3rd order gave solutions for $\nu_1$and ν_2 which were in the complex plane. One out of the four solutions given by simultaneously solving equations (3.9) and (3.10) gave real values, but in return gave out non-real saddlepoint approximations. Using Renshaw's extreme solution of setting the mixed cumulants ($\kappa_{12}(t)$and$\kappa_{21}(t)$) and $\kappa_{22}(t)$ to zero, the solutions are still complex numbers. Solving for $\nu_1$and ν_2 and using those values for the starting values in the iterations in Newton-Raphson's algorithm, the results still don't give real values for the saddlepoint approximation. For this system, the algorithm using the original SAVI values was tested and the same algorithm was applied to the scaled SAVI values. The same steps where repeated on a 4th order cumulants system using the scaled SAVI values, and Mathematica was giving out nine pairs of different solutions for $\nu_1$and ν_2. Again the same problem was encountered; the real values of $\nu_1$and ν_2 gave complex saddlepoint approximations and the complex ν values gave complex values for the saddlepoint.

With these problems encountered using orders ≥ 3, it was decided to restrict the research on the 2nd order case. The Heston model uses its parameters in its option pricing methodology given in section 4, hence using the 2nd order estimated parameters, options can be priced and some comparison with other option pricing methods can be done. This is done in the following section.

7. Option pricing

Now that the parameters have been estimated, it will be useful to create a pseudo-market and price some options on it and compare some of the well known methods to the Heston pricing shown in section 4. There is a lot of literature on pricing options using binomial and trinomial trees, therefore in this section the Heston model's option pricing method is compared to a few tree option pricing models and other recognized methods. These tree models usually differ in the transitional probability calculation and the up and down movements in the trees. Mathematical calculations, equations and derivations won't be shown on this paper, as the particular interest is mainly in the final price of the option. Some parts of the coding used for the methods can be found in the appendix. Some of the methods used for comparisons are the Cox, Ross and Rubinstein (1979) binomial tree which is a discrete-time model for pricing options. Pricing will also be done using the Trigeoris (1991) method and the Hull and White (1993) method. Non-tree methods are also considered; which are the Explicit Finite Difference Method (EFDM), the Fully Implicit algorithm used with Thomas's Algorithm (1995) for option pricing, the Crank and Nicholson (1947) method and also the *MCMC option pricing (abbreviated by MCMC-OptionPriceMeth to avoid confusion with the original MCMC work)* which is explained in detail in van der Voort (2000), which are considered to be numerical methods. The MCMC-OptionPriceMeth mentioned here prices the options as follows (also mentioned in the Appendix):

A risk-neutral world is assumed and the change in the price of the underlying asset is given by the stochastic differential equation $dS = (r - q)Sdt + \sigma\ (S\ ,t)SdW$, where σ is taken as the stochastic volatility, r as the drift and q as the dividend rate of the asset. The SDE can then be integrated as

$$S_{t+\Delta t} = S_t exp\left\{\left(r - q - \tfrac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}\alpha\right\}, \tag{7.1}$$

for time intervals Δt and α being a small change in the Brownian motion. The options priced here are vanilla; with payoff given by the following expression:

$$V_{payoff}(S) = \max(S_T - K, 0), \tag{7.2}$$

where T is the maturiy and K is the strike price. Using (7.1) and (7.2), the price of the derivative can be approximated as:

$$V \approx e^{-rT}\frac{1}{q}\sum_{i=1}^{q} V_{payoff}(S^i)$$

where S^i is simulated by the integration in (7.1) and q is the number of iterations. The code for pricing V is shown in Appendix D.

The above mentioned methods were programmed on Excel VBA except for the Black-Scholes and the Heston, which were evaluated in Mathematica® using free available code from Wolfram®.

7.1 Pseudo Market example

For an example, take the strike price K as 45, the current asset price S as 30, the long-term volatility as 30% and the risk free rate as r = 7% and there are no dividend payments and let the expiry be 6 months and the option be a European Call option. The results are shown in table (7.1). The options priced on trees are taken to have 100 tree steps and price the options using K, S, the long-term volatility, the risk free rate and the expiry dates as stated. The *MCMC Option Pricing* along with the other numeric pricing methods also use the same values mentioned. The *MCMC-OptionPriceMeth* runs only 1000 MCMC iterations to price the option. For pricing the Heston model, the parameters estimated in section 6 are used. The Black-Scholes model also prices the option using the mentioned strike, volatility, stock price and expiry date values. The approximated values of the parameters from the Analysis section will be used after this comparison.

Table 7.1: Comparison of Option values

Method	Option value
Cox-Ross-Rubinstein	0.123702
Trigeoris	0.123693
Jarrow and Rodd	0.123505
Explicit Finite Difference Method (EFDM)	0.124070
Fully Implicit algorithm (Thomas's Algorithm)	0.155557
Crank-Nicholson	0.155417
MCMC-OptionPriceMeth	0.082195
Black Scholes	0.285981
Heston model (In this paper)	0.171935

As seen from the table, the Heston model has not priced far from the other pricing methods. An observable price which is different from the overall observation is the Black-Scholes price which is almost double the average of the other prices.

7.2 Application on the collected market data

In the following results, the estimated values of the parameters from section 6 in table (6.2) are considered. In this market setup the Index levels from the ALSI are considered as the market prices of the asset that is going to be traded, with the SAVI as the volatility. At time t, the price of the asset is $S(t)$and the strike price $K$of the option written is $S(t+\Delta)$. This will only be done on the first 10 values available from the collected data, hence table (7.2) shows 10 option prices obtained from the Heston model and the Black-Scholes model.

Table 7.2: Comparison of Option values between the Black-Scholes and the Heston model

$t+\Delta = 1day$	$\mathbf{S(t)}$	**K**	**Heston model**	**Black-Scholes model**
1	25782.4	25679.2	264.211	263.686
2	25679.2	25605.2	246.073	245.746
3	25605.2	25933.8	75.7986	78.5894
4	25933.8	25915.9	216.695	216.784
5	25915.9	25596.9	412.982	411.38
6	25596.9	25666.8	169.932	170.729
7	25666.8	25820.6	134.245	135.75
8	25820.6	26023.4	116.803	118.711
9	26023.4	26311.4	90.0658	92.6126
10	26311.4	26521.1	117.912	119.88

From table (7.2), it can be seen that the Heston model is not pricing very far from the Black-Scholes model using the estimated parameters.

7.3 Volatility smile graphs of the Heston.

Using the parameter estimates in table (6.2), the implied volatility surface of the Heston model is shown in figure (7.1) below. The stock price is the first observation from the Index Level vector, the strike prices are measured with regards to moneyness and the interest rate is taken to be the approximated stock drift.

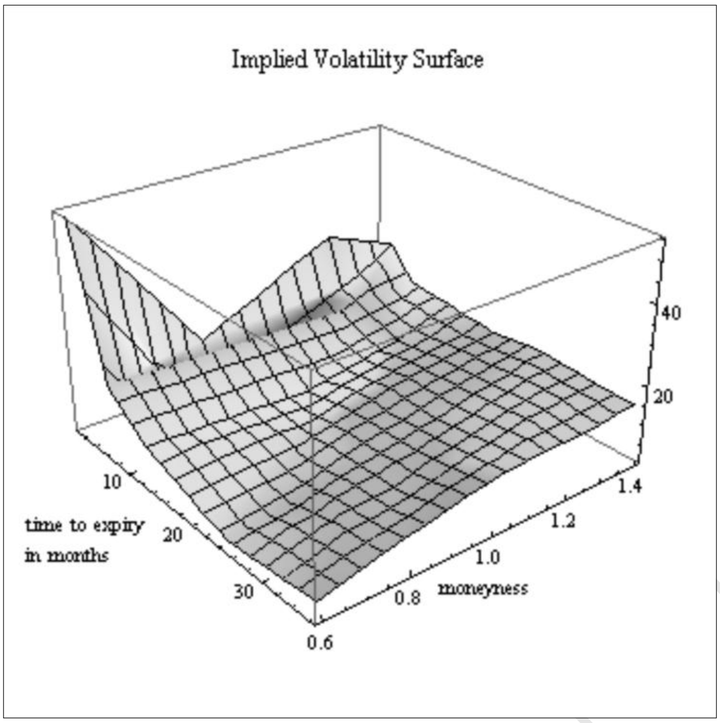

Figure 7.1: Implied volatility surface of the Heston model

<u>*Definition 7.1: Moneyness*</u>

A measure of the degree to which a derivative is likely to have positive monetary value at its expiration, in the risk-neutral measure.

As in most cases (Gkamas, 2010), the surface generated by the Black-Scholes model near time of expiry 0 the skewness is very high and also the same applies for higher moneyness that goes with small expiry times. Looking at figure (7.1), the skewness has decreased near smaller expiry dates even though there is a high pick for small expiry and small moneyness. With higher expiries the surface seems to be flat, including points of higher moneyness.

8. Conclusion

This paper proposed a method of estimating the parameters of the Heston model through the saddlepoint approximation method. The procedure for the method is to obtain a system of ordinary differential equations (ODE) in terms of the cumulants of the model and by running through market data, the cumulants can be obtained by solving these ODEs at discrete time intervals. With the cumulants obtained, they can be substituted in the saddlepoint function.. The saddlepoint approximations are used to estimate the evolution of the probability distribution of the diffusion process, and in doing so the probabilities are in turn used to approximate the likelihood of the model and making it possible to estimate the parameters.

Unlike other methods like the filtering method used by Zirilli (2007), option prices are not needed in the optimisation to solve for the parameters. Also unlike the method used by Renshaw (2000) to estimate the cumulants, the true cumulant values are not needed in this method for the optimisation. This method makes it possible to take a non-reducible diffusion system and simplify it to a system of ODEs that are much easier to integrate and solve in any scientific package.

A disadvantage in the method realised during the research is its failure in higher order cumulant truncations. As seen on this paper, analysis was only done for 2nd order in which the bivariate normal density was used. For 3rd and 4th orders, the method started to be instable and produced complex numbers both in solving for the needed $\theta_1$and θ_2 and the saddlepoint approximations. Its failure sometimes depends on the number iterations ran, hence some other starting values for the parameters or a new jump could be proposed and checked to see how the model behaves. Further research needs to be done on how these issues can be solved so that higher order estimation can be done as it would bring about more accurate estimation of the parameters.

Using the estimated parameters, it is evident that the Heston model prices accurately and close to the other option pricing models. Hence the estimated parameters obtained using the proposed method can be regarded as reliable and can be used in the pricing. The Heston model seems to capture the volatility smile but with the research done, conclusions on whether the Heston model has removed the biasness of the option prices from the Black-Scholes are outside the scope of this paper.

In this paper, the method is applied on available raw market data. Instead of using the SAVI and index levels, the method can be checked in other traded stocks in industries like the commodity prices and the volatility can be determined instantaneously using the stock prices instead of using the SAVI values. In this paper the method is only checked using the Heston model, hence other multivariate diffusion processes can be looked at and also its behaviour in other univariate processes.

9. References

Ait-Sahalia, Y., (2002). *Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach*. Enonometrica 70, 223-262.

Ait-Sahalia, Y., Yu, J., (2006). *Saddlepoint approximation for continuous-time Markov processes*. Journal of Econometrics 134, 507–551.

Ait-Sahalia, Y., Lo, A., (1998). *Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices*.Journal of Finance53, 499-547.

Backus, David, SilverioForesi, and Liuren Wu, (1997). *Accounting for biases in Black-Scholes*, Working paper, New York University.

Black, F. and M. Scholes. (1973). *The pricing of options and corporate liabilities.* The Journal of Political Economy, Vol. 81, No. 3. (May - Jun., 1973), pp. 637-654.

Bondarenko, D. A., Thulasiram, R. K., (2003). *Performance evaluation of parallel algorithms for pricing multidimensional financial derivatives*, Parallel and Distributed Computing Practices.

Brennan, M. J., Schwartz, E. S., (1977). *Convertible bonds: valuation and optimal strategies for call and conversion*,Journal of Finance

Crank, J., and Nicholson, P., (1947). *A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat-Conduction Type*, Proc. Camb. Phil. Soc, Vol. 43, pp. 50-67.

Daniels, H. E., (1954). *Saddlepoint Approximations in Statistics*, Annals of Mathematical Statistics, 25, 631-650.

Duffie, Darrell, Pan, Jun, & Singleton, Kenneth, (2000). *Transform analysis and asset pricing for affine jump-diffusions*.Econometrica, 68(6), 1343–1376.

Eraker, B., (2007). *MCMC Analysis of Diffusion Models with application to Finance*. Journal of Business and Economic Statistics 19 177–191.

Franz, (2010). *Stochastic Volatility Option Pricing using Heston's SV model*, UK Non-linear Time Series Analysis, WS, University of Vienna

Gatheral, J., (2010). *No-dynamic-arbitrage and market impact*, Quantitative Finance 10, 749–759

Gkamas, Dimitrios and D. Paxson, 1999, *Which Stochastic Volatility Model for Option Pricing?*, Working Paper, Manchester Business School

Heston, S., (1993). *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*. The Review of Financial Studies, Vol. 6, No. 2. (1993), pp. 327-343.

Huzurbazar, S., and Huzurbazar, A. V. 1999, *Survival and Hazard Functions for Progressive diseases using Saddlepoint Approximations*, Biometrics, 55, 198-203.

Melino, A., Turnbull, S., (1990). *The pricing of foreign currency options with Stochastic volatility*, Journal of Econometrics, 45: 239-265.

Melino, A., Turnbull, S., (1991). *The pricing of foreign currency*, Canadian Journal of Econometrics, 24: 251-281.

Mikhailov, S. and Nogel, U. (2003). *Heston's stochastic volatility model implementation, calibration and some extensions*, WILMOTT magazine.

Moodley, N., (2005). *The Heston Model: A Practical Approach*. Faculty of Science, University of the Witwatersrand, Johannesburg.

Reid, W.T., (1972), *Riccati differential equations*, Mathematics in Science and Engineering 86, Academic Press, New York.

Renshaw, E., (2000). *Saddlepoint approximations for stochastic processes with truncated cumulant generating functions*. IMA Journal of Mathematics Applied in Medicine & Biology 15, 41-52.

Stein, J.,(1989). *Overreactions in the Options Markets*. Journal of Finance 44:1011-1024.

Stein, E.M, Stein, J.C, (1991). *Stock Price Distribution with Stochastic Volatility: An Analytic Approach*. Review of Financial Studies, 4: 727–757.

Trigeoris, L., (1991), *A Log-transformed binominal numerical analysis method for valuing complex multi-option investments*. Journal Financial and Quantitative Analysis, 26 (3), 309-326.

Turnbull, M. S., Melino, A., (1991). *The Pricing of Foreign Currency Options*, The Canadian Journal of Economics, Vol. 24, No. 2(May, 1991), pp. 251-281 van der Voort, M., (2000). *Implied Modelling*, Erasmus University Rotterdam, Amsterdam, March 2000

Varughese, M., (2010). *Parameter Estimation for a Multivariate, Coupled Diffusion System*. Department of Statistical Science, University of Cape Town, Cape Town.

Varughese, M., Fatti, L., (2008). *Incorporating environmental stochasticity within abiological population model*.Theory.Popul. Biol. 74, 115-129.

White, A., Hull, J., (1987). *The Pricing of Options on Assets with Stochastic Volatilities*, The Journal of Finance, Vol. 42, No. 2 (Jun., 1987), pp. 281-300

Zirilli, F., Mariani, F., Pacelli, G., (2007). *Maximum likelihood of the Heston stochastic volatility model using asset and option prices: an application of non-linear filtering theory*. Optimization Letters (2008) 2:177–222.

Zhylyevskyy, O., (2005). *Pricing American-style derivatives under the Heston Model dynamics: A Fast-Fourier transformation in the Geske-Johnson scheme*.University of Virginia, Charlottesville.

10. Appendix

Appendix A:

This appendix shows the derivation of the bivariate normal distribution used in section (4).

The CGF in equation (2.10) for m = 2 is:

$f(x_1, x_2) \cong \frac{exp\{K(\theta_1,\theta_2)-\theta_1 x_1-\theta_2 x_2\}}{2\pi\sqrt{|K''(\theta_1,\theta_2)|}}$.

Taking $\binom{x_1}{x_2}$ to be normally distributed with mean $\binom{0}{0}$and variance-covariance $\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}$, with correlation$\rho$ and from Renshaw (2000), take $\frac{\partial K}{\partial \theta_i} = x_i$to be true. The CGF becomes:

$K(\theta_1, \theta_2) = \frac{1}{2}(\theta_1^2\sigma_1^2 + 2\rho\sigma_1\sigma_2\theta_1\theta_2 + \theta_2^2\sigma_2^2)$, whence using the above relationship the following is true:

$\frac{\partial K}{\partial \theta_1} = \theta_1\sigma_1^2 + \rho\sigma_1\sigma_2\theta_2 = x,$

$\frac{\partial K}{\partial \theta_2} = \theta_2\sigma_2^2 + \rho\sigma_1\sigma_2\theta_1 = y.$

Solving for θ_1 and θ_2 above yields:

$\theta_1 = \frac{x\sigma_1 - \rho\sigma_1 y}{\sigma_1^2\sigma_2(1-\rho^2)}$,

$\theta_2 = \frac{y\sigma_1 - \rho\sigma_2 x}{\sigma_2^2\sigma_1(1-\rho^2)}$.

Substituting these into the CGF for m = 2 yields the bivariate normal density:

$$f(x, y) = \left[2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}\right]^{-1} exp\left\{\frac{-1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right\}.$$

Appendix B:

This Appendix show the code used to calculate the implied volatility for the Heston model in section (5). The program is copied straight from Mathematica, and it can be found on the page written on the references. In the code, Hhat is the characteristic function used in the evaluation of the complex (Non-Real) integral.

```
Hhat[k_,V_,τ_,θ_,ξ_,ω_,ρ_]:=Module[{f1,f2,d,g,h,ttheta,t,tomega,tc}, c=1/2(k^2 –
™k) ;t=1/2ξ^2 τ;tomega=2/ξ^2 ω;tc=2/ξ^2 c;ttheta=2/ξ^2(™kρξ+θ);
d=√(ttheta² + 4tc);g=1/2(ttheta+d);h=(ttheta+d)/(ttheta-d);f1=tomega(tg - Log[(1-
hExp[dt])/(1-h)]);
f2=g ((1-Exp[d t])/(1-hExp[dt]));Exp[f1+ f2V]];
kim=0.6;Λ =125;

HestonCallLewis[S_,K_,τ_,r_,d_,V0_,MR_,ρ_,Vinf_,VolVol_]:=Module[{Integral,X,θ,ξ,ω,V
,k},X=Log[S/K]+(r-d)τ;k=kre+™kim ;
Integral=(1/π)NIntegrate[Re[Exp[-™k X] (Hhat[k,V0,τ,MR,VolVol,Vinf MR,ρ]/(k2 –
™k))],{kre,0,Λ},Method→{"LobattoKronrodRule","Points"→25},PrecisionGoal→4];
SExp[-dτ] - KExp[-rτ]Integral];

Moneyness={0.6, 0.75, 0.9,1,1.1,1.25,1.4};
Tenors={1,3,6,12,18,24,30,36};
m=Length[Moneyness];t=Length[Tenors]; size=tm;
S=100;r=0.239468;d= 0;
Ncdf=Compile[{{x,_Real}},(Erf[x/√2]+1)/2];

BSCall=Compile[{{S,_Real},{K,_Real},{σ,_Real},{T,_Real},{r,_Real},{d,_Real}},Module[{d1,d2},
d1=(Log[S/K]+(r- d)T+σ^2T/2)/(σ√2);d2=d1–σ√2 ;
S©-^dT Ncdf[d1]-K ©-r T Ncdf[d2]]];

ImpliedVolCall=Compile[{{p,_Real},{S,_Real},{K,_Real},{T,_Real},{r,_Real},{d,_Real},{precision
,_Real}},Module[{Vol0,Vol1,Vol2,Price0,Price1,Price2},Vol2=0.8; Vol0=0;
```

```
Price0=SExp[-dT]-KExp[-rT];If[Price0<0,Price0=0.];
If[p<Price0,0.000001,Vol1=Vol2;Price2=BSCall[S,K,Vol1,T,r,d];
While[Price2<p,Price2=BSCall[S,K,Vol1,T,r,d];Vol1= 2Vol1;];
   Vol2=Vol1;Price1=BSCall[S,K,Vol1,T,r,d];
While[Abs[Price1-p]>precision,Vol1=(Vol0+ Vol2)/2;Price1=BSCall[S,K,Vol1,T,r,d];
If[Price1<p,Vol0=Vol1,Vol2=Vol1]; ]; Vol1]]];
```

Appendix C:

This Appendix shows the Black-Scholes pricing code in Mathematica®.

```
Ncdf=Compile[{{x,_Real}},(Erf[x/√2]+1)/2];
BSCall=Compile[{{S,_Real},{K,_Real},{σ, _Real},{T,_Real},{r,_Real},{d,_Real}},Module[{d1,d2},
d1=(Log[S/K]+ (r-d)T+σ^2T/2)/(σ√2]);d2=d1-σ√2]; S©^-dT(1-PDF[NormalDistribution[0,1],d1])-
K©^-rT(1-PDF[NormalDistribution[0,1],d2])]];
```

Appendix D:

Not all the Excel code used for the option pricings in section 7 will be shown. This Appendix shows the *MCMC-OptionPriceMeth* code for the Option pricing.

```
Private Function MC_Method(S0, r, vol, W(), t)
total_sqd = 0: k = 0: Sum = 0                                  'starting with nothing
deltaT = t/n      '(total time)/(number of iterations)
nudt = (r-((vol ^ 2) / 2)) * deltaT                            'vol = volatility, r = drift
sigdt = vol * Sqr(deltaT)
  For i = 1 To n                                               'generate the stock price and calculate the tolerance
      St = S0
      For j = 1 To 1000
        z = W(i, j)                                            'Draw from Norm distribution
        St = St * Exp(nudt + (sigdt * z))                      'generate stock prices
      Next j
payoff = Application.WorksheetFunction.max(St - X, 0)          'calculating the payoff of the option, X = strike
      Sum = Sum + payoff
total_sqd = total_sqd + (payoff ^ 2)
mean_sqd = (1 / n) * (Sum ^ 2)
      s2_N = ((Exp(-2 * r * t)) / (n - 1)) * (total_sqd - mean_sqd)
      l = s2_N / n
SEofmean = Sqr(l)
tolerance = SEofmean * Application.WorksheetFunction.NormSInv((1 - alpha) / 2)
Tol(1, i) = tolerance
      k = k + 1
      If k = 1000 Then
        epsilon_ = tolerance - 0.001
      End If
  Next i
MC_Method = Sum * (1 / n) * Exp(-r * t)                        'take average to get the MC option value
End Function
```